The PRESAT-vector: Asymmetric T-vector for high-throughput screening of soluble protein domains for structural proteomics

NATSUKO GODA,[1] TAKESHI TENNO,[2] HIROTOSHI TAKASU,[1] HIDEKAZU HIROAKI,[1]
AND MASAHIRO SHIRAKAWA[1]

[1]Graduate School of Integrated Science, Yokohama City University, Yokohama, Kanagawa 230-0045, Japan
[2]Graduate School of Science and Engineering, Ehime University, Matsuyama, Ehime 790–8577, Japan

(RECEIVED September 23, 2003; FINAL REVISION November 11, 2003; ACCEPTED November 23, 2003)

Abstract

A rapid unidirectional method for cloning PCR-amplified cDNA fragments into virtually any fusion protein expression vector is described. The method, termed PRESAT-vector cloning, is based on a T-vector technique that does not require restriction endonuclease digestion of the PCR product. Subsequently, we applied a novel ORF selection method of the ligated plasmid products. This second step involves restriction endonuclease treatment that eliminates the plasmids containing an ORF in the wrong orientation prior to transformation into the bacterial host for further protein expression studies. To achieve this selection, we customized the 5′-sequence of the "rear" PCR primer corresponding to the C terminus of the protein to be expressed. The colonies harbored only the ligated products of the desired orientation at >90% efficiency. This method is applied to a GST fusion expression system, and an HTS system for soluble proteins from an expression library was tested.

Keywords: structural proteomics; recombinant fusion protein; T-vector cloning; asymmetric T-vector; unidirectional PCR cloning; high-throughput screening; glutathione-S-transferase

Supplemental material: See www.proteinscience.org

With the completion of a growing number of genome sequencing projects (Wakeland and Wandstrat 2002; also see http://www.ncbi.nlm.nih.gov/), the next major challenge is to focus on the genome-wide analysis of protein structure/function relationships (Yokoyama et al. 2000; Burley and Bonanno 2002; Phizicky et al. 2003). The challenge of studying proteins on a genome-wide scale requires the development of high-throughput approaches in protein expression, purification, biochemical analysis, and structure determination. Cloning and expression in *Escherichia coli* are favored in terms of ease of handling, rapid growth rate, and reasonable costs. However, one major disadvantage of overexpression of heterologous proteins in *E. coli* is that many proteins tend to be expressed as inclusion bodies. In order to carry out these studies, an HTS system is required that can rapidly generate plasmid constructs and find successful clones that express soluble proteins and/or protein domains.

A T-vector cloning technique is one of the most efficient methods for subcloning of PCR products (Zhou and Gomez-Sanchez 2000). In structural and/or functional proteomics, the PCR-based method has become routine for the parallel preparation of large numbers of cDNA fragments. However,

---

Reprint requests to: Hidekazu Hiroaki, Division of Molecular Biophysics, Graduate School of Integrated Science, Yokohama City University, 1-7-29 Suehirocho, Tsurumi, Yokohama, Kanagawa 230-0045, Japan; e-mail: hiroakih@tsurumi.yokohama-cu.ac.jp; fax: 81-45-508-7361.

Abbreviations: NMR, nuclear magnetic resonance; GSH, glutathione; GST, glutathione-S-transferase; HTS, high-throughput screening; IPTG, isopropyl-D-thiogalactopyranoside; LB, Luria Broth media; HSQC, heteronuclear single quantum coherence; OD, optical density; SDS-PAGE; sodium dodecylsulfate-polyacrylamide gel electrophoresis; CDNB, 1-chloro-2,4-dinitrobenzene; PCR, polymerase chain reaction; ORF, open reading frame.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.03439004.

Copyright © 2004 The Protein Society

it is well known that these PCR products are unusually resistant to efficient cloning, especially using either standard sticky-end (Kaufman and Evans 1990) or blunt-end cloning techniques. Thus many laboratories have chosen to use the T-vector cloning methodology when starting their research projects.

However, a major disadvantage of the T-vector cloning method for the construction of expression libraries (either direct expression or fusion-tagged vectors) is that the cloning is not directional. Because the T-vector technique is based on the formation of a single base pair between the overhanging 3′-T and 3′-A, the population of preligated intermediates of vector-insert complexes is equal for both orientations of a PCR product. Thus, we developed a selection method to eliminate ligated products containing an ORF in an undesired orientation. To achieve this, we designed a locally asymmetric sequence at the T-site of the vector.

Two distinct methods for preparing the single 3′-T overhang at the cloning site of the vector are reported. The first method involves linearizing a vector with a blunt-end restriction enzyme, such as EcoRV, and a single thymidine is subsequently added at the 3′ end by Taq polymerase with an excess amount of dTTP (Holton and Graham 1991; Marchuk et al. 1991; Papp et al. 1995). Alternatively, a pair of 3′ single dT overhangs is generated in a T-vector by digestion using a suitable restriction endonuclease. For this purpose, XcmI (Cha et al. 1993; Kwak and Kim 1995; Borovkov and Rivkin 1997; de Vries 1998) and AspEI/Eam1105I/AhdI (Ichihara and Kurosawa 1993; Ido and Hayami 1997; Jeung et al. 2002) were used. The latter enzymes are isoschizomers that generate a 3′-overhanging site at the sequence 5′-GATNNN/NNATC-3′, where N represents any of the four bases. We introduced a pair of AhdI sites in the T-vector, which allowed us to design asymmetric sequences adjacent to the T-overhanging sites.

In this study, we propose a novel TA-cloning-based methodology for unidirectional cloning of PCR-amplified cDNA fragments into fusion protein expression vectors, such as a GST expression system (Smith and Johnson 1988). In the designed asymmetric TA-cloning sites, we engineered a potential NcoI/NdeI site only within the upstream position of the pair of AhdI recognition sites, but not at the downstream position (Fig. 1A), thereby creating asymmetric TA-cloning sites. This engineered second restriction enzyme site is further used for the selection of ORF orientation, in combination with appropriately designed PCR primer sequences and simple restriction enzyme treatment before transformation. The efficiency of selectivity, feasibility, and applicability of the method is also discussed.

Results and Discussion

Figure 1 shows the schematic representation of the selection methodology of ORF orientation of an insert cDNA fragment using a simple restriction enzyme treatment. To distinguish only the plasmid of the forward (desired) ORF orientation from that of the reverse (undesired) orientation, we used an inexpensive screen by restriction enzyme digestion selection prior to transformation of the plasmid into the bacterial hosts. To achieve this, we designed the 5′-terminal nucleotides of the PCR primers as follows: the rear primer of the PCR must start with 5′-GG or 5′-ATG sequences, and the forward primer must *not* start with 5′-GG or 5′-ATG, respectively. Because we engineered a potential NcoI/NdeI site upstream of the TA-cloning site, a new NcoI site will appear only when the PCR product is ligated into the TA-cloning site in a reverse orientation, because the rear PCR primer contains the 5′-GG sequence. Likewise, when the rear primer starting with the 5′-ATG sequence is cloned into the TA-cloning site of the vector, a new NdeI site instead of the NcoI site will appear at the same place. Thus, the undesired ligation products, which have a reverse orientation of insert DNA, are linearized by subsequent digestion with either NcoI or NdeI. The linearized plasmids were then eliminated during the subsequent transformation step (Fig. 1B). We named this novel method "PRESAT-vector" (Potential Restriction Enzyme Selectable Asymmetric T-vector).

Figure 2 shows the agarose and SDS-PAGE gel of the parallel construction of GST fusion expression vectors. To demonstrate the efficiency of our PRESAT-vector method, we converted a pGEX-4T3-derived GST fusion protein expression vector to the T-vector, pGEX-4T3-PRESAT. Five genes encoding a putative protein domain of ~70 amino acids, named "MIT" (Ciccarelli et al. 2003), were used for the demonstration. The genes were PCR-amplified from yeast genomic DNA, or from commercially available mouse and human first-strand cDNAs using standard Taq polymerase. Aliquots of PCR products were ligated with the linearized pGEX-4T3-PRESAT. *E. coli* DH5alpha was transformed with the ligation mixture, and plasmid DNA consisting of a mixture of the forward and reverse ORF was prepared from liquid culture. The procedure for selecting constructs containing insert DNA in the desired orientation was then applied. The plasmids were digested with NcoI, and subsequently transformed into BL21(DE3). Five colonies from each of the five constructs were isolated and used for expression experiments. A total of 23 of 25 colonies (i.e., 92%) were shown to express fusion proteins of the expected molecular mass (32–34 kD) by SDS-PAGE analysis (Fig. 2A). Each colony expressing the GST-fusion protein of the expected size was further analyzed for protein solubility. The whole bacterial lysate, and soluble and pellet fraction of the sonicated lysate, were analyzed by both SDS-PAGE and CDNB assay (Fig. 2B,C). Independently, five colonies from all five constructs were checked by colony PCR screening to estimate the ratio of ligated product without insert (Fig. 2D). A total of 23 of 25 colonies (i.e., 92%)

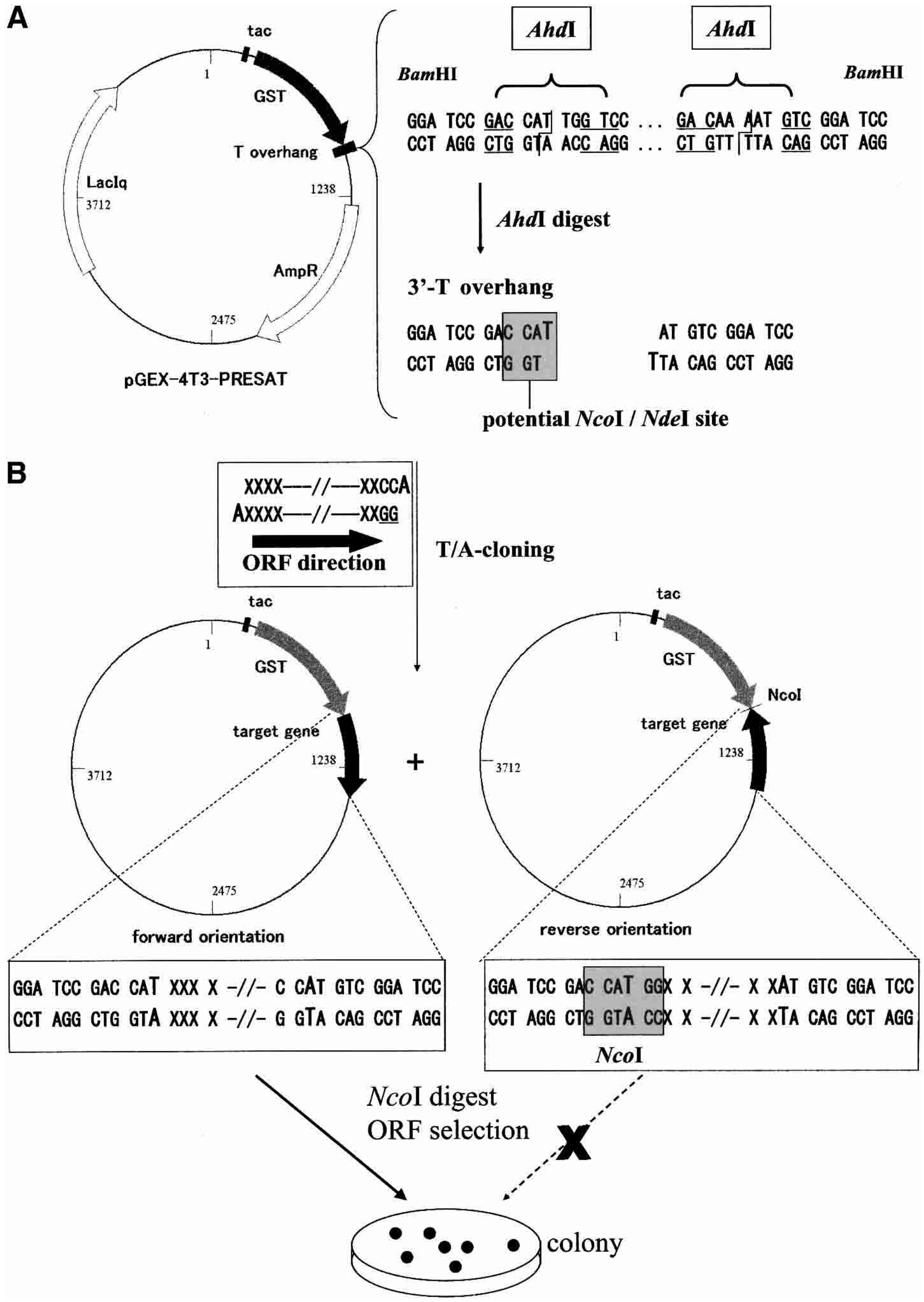
A
tac
1
GST
T overhang
1238
LacIq
3712
AmpR
2475
pGEX-4T3-PRESAT
AhdI
AhdI
BamHI
BamHI
GGA TCC GAC CAT TGG TCC ... GA CAA AAT GTC GGA TCC
CCT AGG CTG GTA ACC AGG ... CT GTT TTA CAG CCT AGG
AhdI digest
3'-T overhang
GGA TCC GAC CAT
CCT AGG CTG GT
AT GTC GGA TCC
TTA CAG CCT AGG
potential NcoI / NdeI site
B
XXXX—//—XXCCA
AXXXX—//—XXGG
ORF direction
T/A-cloning
tac
1
GST
target gene
1238
3712
2475
forward orientation
+
tac
1
GST
NcoI
target gene
1238
3712
2475
reverse orientation
GGA TCC GAC CAT XXX X -//- C CAT GTC GGA TCC
CCT AGG CTG GTA XXX X -//- G GTA CAG CCT AGG
GGA TCC GAC CAT GGX X -//- X XAT GTC GGA TCC
CCT AGG CTG GTA CCX X -//- X XTA CAG CCT AGG
NcoI
NcoI digest
ORF selection
X
colony

Figure 1. Concept of the asymmetric directional T-vector. (*A*) Construction of pGEX-4T3-PRESAT and direct cloning of PCR product in pGEX-4T3-PRESAT. (*B*) The schematic representation of the ORF selection method using potential restriction enzyme site. The figure illustrates the case in which NcoI is chosen as the second restriction enzyme for selection. The rear PCR primer is designed with 5′-GG at the 5′ end, so that only the ligated plasmid with insert in the reverse orientation will have the NcoI site at the TA-cloning position. For NdeI selection, the rear primer with 5′-ATG is used instead of the 5′-GG primer.

gave PCR products of the expected length. Therefore, using our PRESAT technology, it is possible to select and to keep only those plasmid constructs containing insert DNA in the forward direction, which express the soluble fusion protein. Thus, the PRESAT strategy can significantly reduce the needless analysis of nonproductive clones, thereby saving both cost and time.

For production of isotopically labeled GST-MIT fusion protein, one of the clones harboring pGST-PRESAT-MIT(mVps4b) was cultured in M9 minimal media with

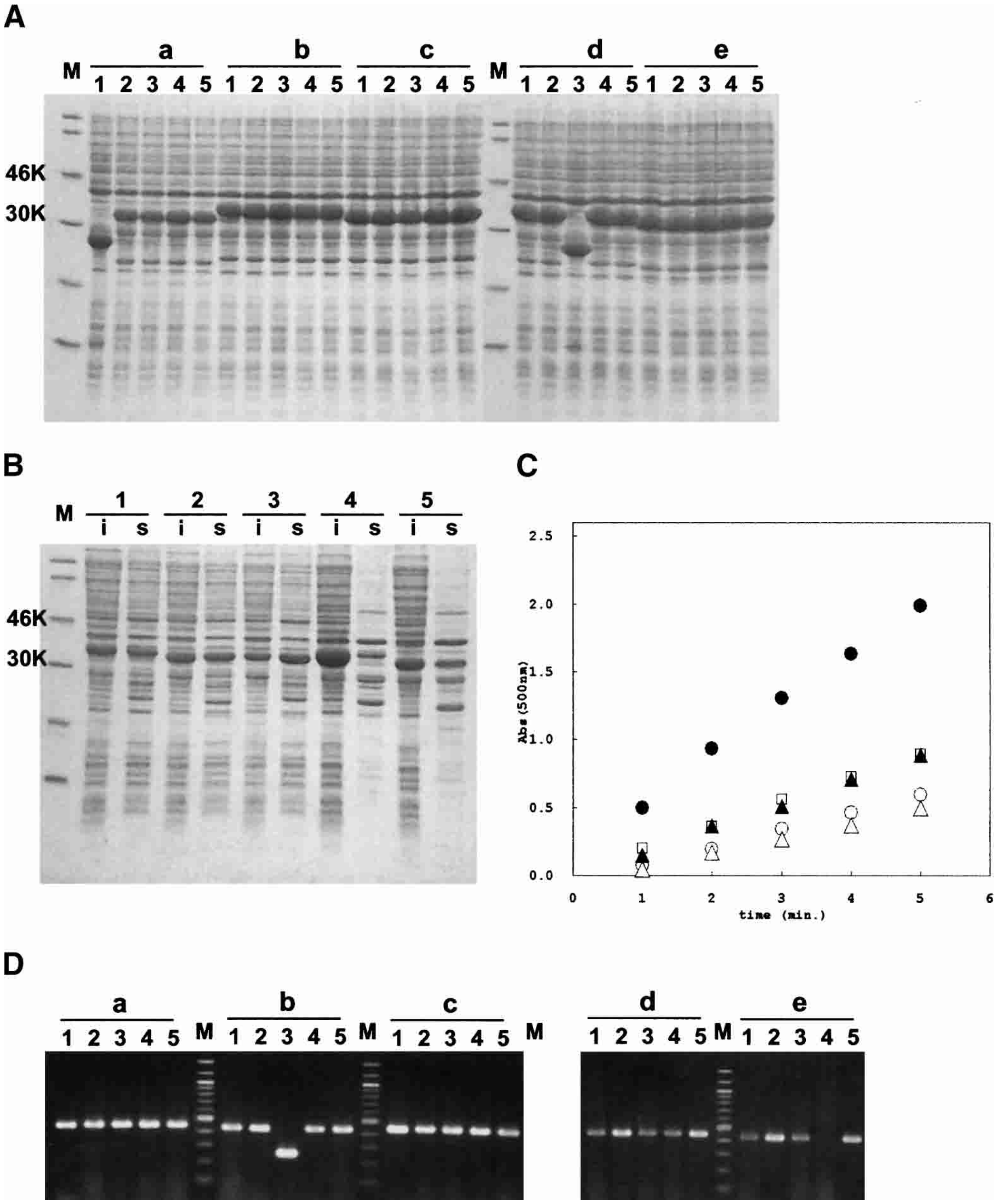

Figure 2. Parallel construction of the five putative protein domains as GST-fusion proteins. (*A*) Overexpression of GST-fusion proteins. Whole-cell extract of randomly selected *E. coli* BL21(DE3) colonies carrying recombinant plasmid were analyzed on 15% SDS-PAGE. (Lane *M*) Molecular standard, (lanes *a1–a5*) yeast Vps4, (lanes *b1–b5*) human Vps4b, (lanes *c1–c5*) mouse Vps4b, (lanes *d1–d5*) human Snx15a, (lanes *e1–e5*) mouse Snx15a. (*B*) Analysis of soluble and insoluble fractions of the GST-MIT fusion proteins. s and i represent soluble and insoluble fractions, respectively. (Lane *1*) Yeast Vps4, (lane *2*) human Vps4b, (lane *3*) mouse Vps4b, (lane *4*) human Snx15a, (lane *5*) mouse Snx15a. (*C*) Colorimetric CDNB assay of soluble fractions of the five GST-MIT fusion proteins. (Open circle) Human Vps4b, (filled circle) mouse Vps4b, (open triangle) yeast Vps4, (open square) human Snx15a, (filled triangle) mouse Snx15a. (*D*) Colony PCR analysis of the randomly selected colonies carrying GST-MIT fusion vector. (Lanes *a1–a5*) Human Vps4b, (lanes *b1–b5*) human Snx15a, (lanes *c1–c5*) yeast Vps4, (lanes *d1–d5*) mouse Vps4b, (lanes *e1–e5*) mouse Snx15a.

^{15}N-NH$_4$Cl as the sole nitrogen source. The fusion protein was purified by GSH affinity chromatography and cleaved by thrombin, and the domain of interest was separated from GST tag using anion-exchanging chromatography. The efficiency of labeling the MIT domain of mVps4b was demonstrated in ^{1}H-^{15}N HSQC spectra (Fig. S1 of Supplemental Material). More than 90% of backbone amide signals were observed as nice sharp peaks in good spectral dispersion, indicating that the MIT domain is a structured monomeric domain. Further structural study of related domains are under investigation.

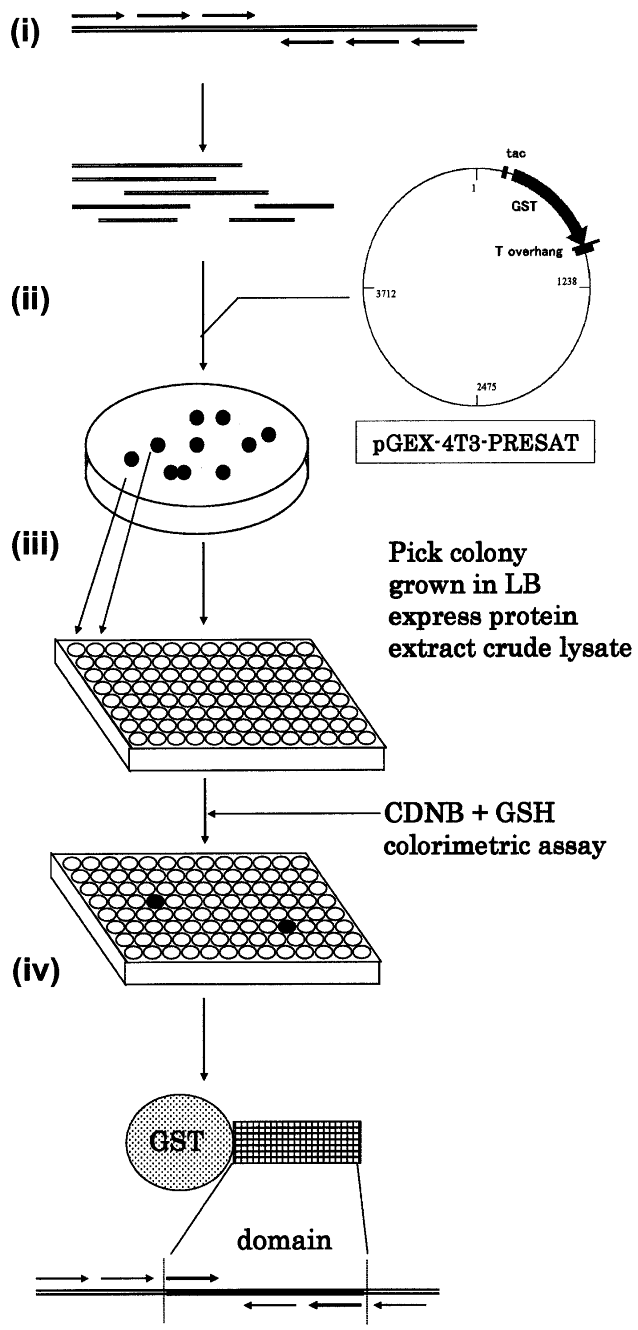

Figure 3. Concept of HTS for the soluble GST-fusion proteins from protein expression libraries. (i) Various PCR primers were designed and used to amplify a series of partial cDNA fragments of the protein of interest, which may contain a putative domain. (ii) The fragments are inserted into pGEX-4T3-PRESAT vector, selected for ORF orientation by restriction cleavage, and transformed into *E. coli* for colony isolation. (iii) Deep 96-well format plates are used to culture individual clones to allow expression of the GST-fusion proteins. (iv) The clone displaying the highest CDNB assay signal will be selected for further analysis.

The selection methodology of the PRESAT-vector system is generally applicable, because it can be adapted to any expression vector system of any host strain by conversion into an asymmetric T-vector using a linker containing the double AhdI sites. In this study we chose to use the *E. coli* GST-fusion expression system (Smith and Johnson 1988), because of ease of detection of the recombinant fusion protein in the crude bacterial lysate. The amount of GST-fusion protein in the soluble cell extract could be readily estimated using the colorimetric CDNB assay. Although the estimation of soluble fraction of the GST-fusion proteins by CDNB was not exactly correlating with the one by SDS-PAGE (Fig. 2B,C), the colorimetric assay is still beneficial for HTS using a 96-well formatted plate. HTS to select soluble fusion proteins using a GST-based protein expression library could result in a large number of false positives arising from ligated pGEX plasmid with an insert of reverse orientation, if conventional T-vector methodology had been applied. However, in our PRESAT-vector system, only the colonies harboring plasmid-encoding soluble GST-fusion proteins, but not GST alone, were isolated (>90% efficiency). The ligation background of TA-cloning is known to be remarkably low, and this feature encouraged us to use the PRESAT-vector system for an HTS to obtain soluble truncated forms of the protein of interest (Fig. 3). The method described in this paper can be easily adapted to laboratory automation using a 96-well plate format and appropriate robotic liquid handling. Some ideas for obtaining soluble fusion protein using colorimetric HTS methods in vivo and in vitro were proposed, including use of green fluorescent protein (Waldo et al. 1999; Kawasaki and Inagaki 2001) or β-galactosidase (Wigley et al. 2001) as a fusion partner. However, GST-fusion protein is still more beneficial, because the affinity purification using GSH beads is easy, the biochemical in vitro ligand binding study is feasible, and the expression level in *E. coli* is sufficiently high for structural biology such as isotopically labeled NMR sample preparation.

In conclusion, we have demonstrated the feasibility and applicability of a PRESAT-vector PCR cloning methodology for use in a recombinant fusion protein expression system, with high efficiency of ORF orientation selection and drastic suppression of cost and time for vector construction. Our PRESAT-vector system has enormous potential for high-throughput structural and functional proteomics studies.

Materials and methods

Restriction enzymes NcoI, NdeI, AhdI (New England Biolabs), and BamHI (TOYOBO) were used. T4 DNA ligase, Wizard plasmid miniprep kit, and Wizard SV DNA gel purification kit (Promega) were used for ligation and plasmid purification. Oligonucleotide primers were obtained from Hokkaido System Science.

pGEM-T (Promega), pGEX-4T3 (Amersham Bioscience), and pET32a (Novagen), respectively, were purchased. For PCR cloning, QUICK-Clone cDNA of mouse spleen and human HeLa cell (Clontech) were used as a template.

Construction of pGEX-4T3-based asymmetric T-vector

Prior to insertion of the linker producing the 3′ T-overhang site, the AhdI restriction site internal to the native AmpR gene was eliminated by introduction of a silent mutation. The 5′ phosphorylated primer dGTTATCTACACCACGGGGAGCCAGGCAACT ATGG was used according to the standard site-directed mutagenesis protocol (Zoller and Smith 1987). The primers of dGGATCC GACCATTGGTCCACCTGACTGACGACAGTTTTGAC [front] and dGGATCCGACATTTGGTCTTATTAGCCCAGCACCGCCG [rear] were used to PCR-amplify an ~550-bp DNA fragment from pET32aPACAP as a template. Note that the 550-bp DNA encoding thioredoxin was cleaved off from the linearized PRESAT-vector. The length of this 550-bp DNA was chosen for complete separation of the linearized PRESAT-vector from the uncleaved vector on agarose gel. This PCR fragment possesses two BamHI sites followed by the AhdI sites. The fragment was first subcloned into pGEM-T vector according to the manufacturer's instruction. The DNA sequence was checked before digestion to obtain a 550-bp BamHI linker fragment. The linker was then ligated into the unique BamHI site of the modified pGEX-4T3, downstream of the thrombin cleavage site. The plasmid with the insert of the desired linker orientation following GST was selected. The vector (2.5 µg) was linearized with AhdI (25 U, 37°C for 60 min) to generate the 3′-single overhanging dT sites. Approximately 1.5 µg of the linearized pGEX-4T3-PRESAT vector (~5 kbp) was obtained after gel purification. The product, pGEX-4T3-PRESAT, was used for further cloning experiments.

Parallel construction of GST-fusion vector expressing a putative protein domain

Appropriate PCR primers to amplify five MIT domains of SNX15a (human and mouse) and Vps4b (yeast, mouse, and human) were designed. The front primers must start with the first codon of any amino acid in order to match the frame with the preceding GST, but the 5′ bases must *not* start with 5′-GG. The rear primers must contain a stop codon after the putative C-terminal boundary of the domains, and 5′-terminal bases must start with 5′-GG. These bases form the second NcoI site only when the insert is ligated to give the reverse ORF orientation to pGEX-4T3-PRESAT. This newly created NcoI site was used for the subsequent selection step. A standard PCR was performed using Taq polymerase (Invitrogen), using a TP240 thermalcycler (TAKARA) at a 25-µL scale. Two-microliter aliquots were ligated with 100 ng of linearized pGEX-4T3-PRESAT at 4°C for 16 h. The ligation mixture was transformed into *E. coli* DH5alpha, and the bacteria were grown at 37°C in 2 mL of LB media containing 50 µg/mL ampicillin for recovery of the plasmids. Aliquots (0.2 µg) of plasmid were treated with 10 units of NcoI for 1 h at 37°C. Finally, 10 ng of the digest was transformed into the expression strain *E. coli* BL21(DE3), and selected on solid medium containing 50 µg/mL ampicillin. Isolated single colonies were analyzed in the protein expression studies.

Expression studies of the GST-fusion proteins

Five clones from each of the five GST-MIT fusion vectors were grown in LB-glucose medium at 37°C. Expression of the recombinant gene was induced with IPTG (1 mM) when the cell density reached an OD of 0.4. Cells were harvested 3 h after induction, and the cell extracts were analyzed by SDS-PAGE using a 15% gel. The cells were broken by sonication in a buffer containing 0.15 M NaCl and 20 mM Tris-HCl (pH 7.5), centrifuged, and analyzed by SDS-PAGE. The quantity of GST-MIT fusion protein in the soluble and insoluble fraction was estimated. The CDNB colorimetric assay for estimation of the soluble fraction of GST-MIT fusion proteins was done according to the manufacturer's instruction. In brief, 2 µL of the sonicated supernatants was added to 1 mL of CDNB assay buffer containing 1 mM CDNB, 1 mM GSH, and 0.1 M potassium phosphate buffer (pH 7.5) at room temperature, and an increase of absorbance at 340 nm was measured.

Preparation of ^{15}N-labeled MIT domain of mouse Vps4b

For ^{15}N labeling of the fusion protein, GST-MIT(mVps4b), a starter culture of BL21(DE3) harboring the plasmid pGEX-PRESAT-MIT(mVps4b), was grown in 2 mL LB containing 50 µg/mL ampicillin at 37°C until the growth was saturated. A 0.1-L culture of M9 minimal media containing 0.7 g/L ^{15}N-NH$_4$Cl as the sole nitrogen source and 4.5 g of glucose was inoculated with 3 mL of the starter culture, and allowed to grow at 30°C for 16 h, and then inoculated into 0.9 L of the same ^{15}N-enriched M9 media. Expression of the recombinant gene was induced with IPTG (1 mM) when the cell density reached an OD of 0.4. Cells were harvested 6 h after induction. The cells were pelleted, resuspended in 50 mL of lysis buffer, broken by sonication, and centrifuged to remove insoluble cellular debris. Supernatant was passed to glutathione-Sepharose Fast Flow column (Amersham Bioscience), and GST-MIT fusion protein was purified by affinity chromatography according to the manufacturer's instruction. Then the fusion protein was treated by bovine thrombin (Amersham Bioscience) at 25°C for 20 h. One hundred fifty units of the enzyme were used against fusion protein from 1 L culture media. The MIT domain was further purified by anion exchanging chromatography as an approximately single band on SDS-PAGE. Typically, 6 mg of purified MIT domain was obtained from 1 L culture of ^{15}N-enriched M9 media.

NMR spectroscopy

NMR experiments were performed on a Bruker Avance DRX 500-MHz NMR spectrometer equipped with a cryogenic triple resonance probe and Z-axis pulsed filed gradients. The NMR sample contained 1.3 mg of ^{15}N-MIT(mVps4b) dissolved in 300 µL NMR buffer containing 50 mM NaCl, 25 mM sodium phosphate buffer (pH 5.5), and 2% glycerol. ^{1}H-^{15}N HSQC spectra (Neri et al. 1989) was acquired with eight transients and 256 increments according to echo-antiecho method (Palmer et al. 1991; Kay et al. 1992) at 25°C, and processed with nmrPipe (Delaglio et al. 1995).

Electronic supplemental material

An example of ^{1}H-^{15}N HSQC spectra of a purified MIT domain of mouse Vps4b obtained by this method is available.

Acknowledgments

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby

marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Borovkov, A.Y. and Rivkin, M.I. 1997. XcmI-containing vector for direct cloning of PCR products. *Biotechniques* **22:** 812–814.

Burley, S.K. and Bonanno, J.B. 2002. Structural genomics of proteins from conserved biochemical pathways and processes. *Curr. Opin. Struct. Biol.* **12:** 383–391.

Cha, J., Bishai, W., and Chandrasegaran, S. 1993. New vectors for direct cloning of PCR products. *Gene* **136:** 369–370.

Ciccarelli, F.D., Proukakis, C., Patel, H., Cross, H., Azam, S., Patton, M.A., Bork, P., and Crosby, A.H. 2003. The identification of a conserved domain in both spartin and spastin, mutated in hereditary spastic paraplegia. *Genomics* **81:** 437–441.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

de Vries, E. 1998. pUCPCR1. A vector for direct cloning of PCR products in a double Xcm1 restriction site offering compatible single 3′-overhanging T residues. *Mol. Biotechnol.* **10:** 273–274.

Holton, T.A. and Graham, M.W. 1991. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res.* **19:** 1156.

Ichihara, Y. and Kurosawa, Y. 1993. Construction of new T vectors for direct cloning of PCR products. *Gene* **130:** 153–154.

Ido, E. and Hayami, M. 1997. Construction of T-tailed vectors derived from a pUC plasmid: A rapid system for direct cloning of unmodified PCR products. *Biosci. Biotechnol. Biochem.* **61:** 1766–1767.

Jeung, J.U., Cho, S.K., Shim, K.S., Ok, S.H., Lim, D.S., and Shin, J.S. 2002. Construction of two pGEM-7Zf(+) phagemid T-tail vectors using AhdI-restriction endonuclease sites for direct cloning of PCR products. *Plasmid* **48:** 160–163.

Kaufman, D.L. and Evans, G.A. 1990. Restriction endonuclease cleavage at the termini of PCR products. *Biotechniques* **9:** 304–306.

Kawasaki, M. and Inagaki, F. 2001. Random PCR-based screening for soluble domains using green fluorescent protein. *Biochem. Biophys. Res. Commun.* **280:** 842–844.

Kay, L.E., Keifer, P., and Saarinen, T. 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114:** 10663–10665.

Kwak, J.H. and Kim, M.Y. 1995. Construction of T-vector for direct cloning and expression of cloned genes in *Escherichia coli*. *Anal. Biochem.* **228:** 178–180.

Marchuk, D., Drumm, M., Saulino, A., and Collins, F.S. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19:** 1154.

Neri, D., Szyperski, T., Otting, G., Senn, H., and Wuthrich, K. 1989. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* **28:** 7510–7516.

Palmer, A.G., Cavanagh, J., Wright, P.E., and Rance, M. 1991. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation spectroscopy. *J. Magn. Reson.* **93:** 151–170.

Papp, T., Kirchner, S., Diener, U., Jafari, M., Golka, A., and Schiffman, D. 1995. Cloning of PCR fragments with a modified M13mp18 T-vector. *Trends Genet.* **11:** 169.

Phizicky, E., Bastiaens, P.I., Zhu, H., Snyder, M., and Fields, S. 2003. Protein analysis on a proteomic scale. *Nature* **422:** 208–215.

Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67:** 31–40.

Wakeland, E.K. and Wandstrat, A.E. 2002. Analyzing genomes: Current realities and future possibilities. *Curr. Opin. Immunol.* **14:** 622–626.

Waldo, G.S., Standish, B.M., Berendzen, J., and Terwilliger, T.C. 1999. Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* **17:** 691–695.

Wigley, W.C., Stidham, R.D., Smith, N.M., Hunt, J.F., and Thomas, P.J. 2001. Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. *Nat. Biotechnol.* **19:** 131–136.

Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., et al. 2000. Structural genomics projects in Japan. *Nat. Struct. Biol.***7 (Suppl.):** 943–945.

Zhou, M.Y. and Gomez-Sanchez, C.E. 2000. Universal TA cloning. *Curr. Issues Mol. Biol.* **2:** 1–7.

Zoller, M.J. and Smith, M. 1987. Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded DNA template. *Methods Enzymol.* **154:** 329–350.